Weekly exercise: Convolutional Neural Network

IN5400 / IN9400 MACHINE LEARNING FOR IMAGE ANALYSIS

Task1: Calculate the theoretical receptive field

You shall implement the function "receptive_field" and use it to calculate the receptive field for 5 different convolutional neural network architectures. When evaluating the receptive field, the image size (resolution) and the size of the objects of interest is important to consider.

- Receptive field: R
- Filter size: F
- Stride: S
- Layer index: k
- $R^0 = 1$ Receptive field of input data

Note: Superscript indicate layer index not exponent

Equation (1)

$$R^k = R^{k-1} + \left[(F^k - 1) \cdot \prod_{i=1}^{k-1} S^i\right]$$

	Architecture 1		Architecture 2		Architecture 3		Architecture 4		Architecture 5	
	FilterSize	Stride	FilterSize	Stride	FilterSize	Stride	FilterSize	Stride	FilterSize	Stride
Layer 1	3	1	3	2	3	2	5	1	5	2
Layer 2	3	1	3	1	3	2	5	1	5	1
Layer 3	3	1	3	2	3	2	5	1	5	2
Layer 4	3	1	3	1	3	2	5	1	5	1
Layer 5	3	1	3	2	3	2	5	1	5	2
Layer 6	3	1	3	1	3	2	5	1	5	1

In [5]:

```
def receptive_field(f, s):
    # Implement equation(1)

    # Inputs:
    # f (list): Filter size for each layer
    # s (list): Stride for each layer

    # Output
    # R: The calculated receptive field for each layer as a numpy array

    # ToDo:
    R = [1]
    for kk in range(len(s)):
        S = 1
        for ii in range(kk):
            S = S * s[ii]
        fov = R[-1] + (f[kk] - 1) * S
        R.append(fov)
    return np.array(R)
```

In [9]:

```
# Defining the architectures

# Architecture1
A1_filterSize = [3, 3, 3, 3, 3, 3]
A1_stride     = [1, 1, 1, 1, 1, 1]
A1_Recept     = receptive_field(A1_filterSize, A1_stride)

# Architecture2
A2_filterSize = [3, 3, 3, 3, 3, 3]
A2_stride     = [2, 1, 2, 1, 2, 1]
A2_Recept     = receptive_field(A2_filterSize, A2_stride)

# Architecture3
A3_filterSize = [3, 3, 3, 3, 3, 3]
A3_stride     = [2, 2, 2, 2, 2, 2]
A3_Recept     = receptive_field(A3_filterSize, A3_stride)

# Architecture4
A4_filterSize = [5, 5, 5, 5, 5, 5]
A4_stride     = [1, 1, 1, 1, 1, 1]
A4_Recept     = receptive_field(A4_filterSize, A4_stride)

# Architecture5
A5_filterSize = [5, 5, 5, 5, 5, 5]
A5_stride     = [2, 1, 2, 1, 2, 1]
A5_Recept     = receptive_field(A5_filterSize, A5_stride)
```

```
[  1   3   7  15  31  63 127]
```

In [7]:

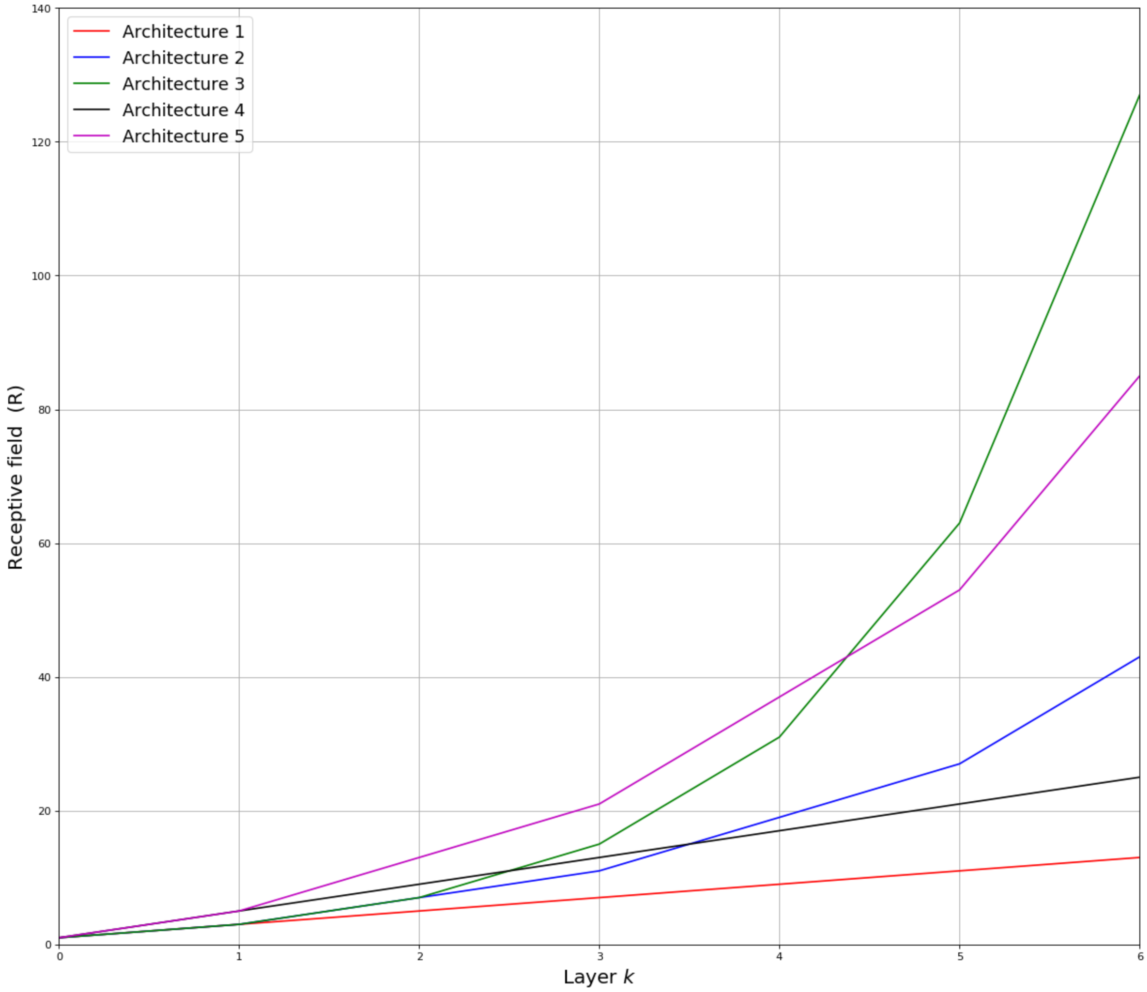
```
import numpy as np
import matplotlib.pyplot as plt
import time

%matplotlib inline

plt.figure(figsize=(18, 16), dpi= 80, facecolor='w', edgecolor='k')
ax = plt.subplot(1, 1, 1)
plt.plot(A1_Recept, 'r', label='Architecture 1')
plt.plot(A2_Recept, 'b', label='Architecture 2')
plt.plot(A3_Recept, 'g', label='Architecture 3')
plt.plot(A4_Recept, 'k', label='Architecture 4')
plt.plot(A5_Recept, 'm', label='Architecture 5')
plt.ylabel('Receptive field  (R)', fontsize=18)
plt.xlabel('Layer $k$', fontsize=18)
ax.grid()
plt.ylim([0, 140])
plt.xlim([0, 6])
ax.legend(loc='upper left', fontsize=16)
```

Out[7]:

```
<matplotlib.legend.Legend at 0x1bc7616ad30>
```



Task2: Convolution

You are given an input image (x), kernel (w) and bias (b). Your task is to evaluate the shaded pixel in the image after the convolution. The origin of the kernel is the shaded pixel.

$$x = \begin{array}{|c|c|c|c|c|}\hline -2 & 4 & 5 & -4 & 1 \\ \hline 2 & 1 & 7 & 5 & 6 \\ \hline 7 & 3 & 2 & -9 & 7 \\ \hline 8 & 2 & 3 & 4 & 9 \\ \hline 6 & -1 & 0 & 1 & 6 \\ \hline \end{array} \qquad w = \begin{array}{|c|c|c|}\hline 2 & 5 & -6 \\ \hline 0 & 2 & 4 \\ \hline 5 & 1 & 3 \\ \hline \end{array} \qquad b = \boxed{2}$$

Solution:

Assume we use zero padding.

$$\begin{aligned} z = & -9 \cdot 2 + 7 \cdot 5 + 0 \cdot -6 \\ + & \quad 4 \cdot 0 + 9 \cdot 2 + 0 \cdot 4 \\ + & \quad 1 \cdot 5 + 6 \cdot 1 + 0 \cdot 3 \\ + & \quad 2 \\ = & \ 48 \end{aligned}$$

Task3: Channel dimention of activation maps

What is the relations between the channel dimention in an activation map and the filter bank?

Solution:

Filter bank: $w.shape = [F_N, F_c, F_h, F_w]$

Activation map: $a.shape = [N_c, N_h, N_w]$

$N_c^i = F_c^i$

$N_c^{i+1} = F_N^i$

Task4: Spatial size of the activation map

Given an activation map with shape $[N_c = 128, N_h = 225, N_w = 225]$ and a kernal with shape $[F_c = 128, F_h = 5, F_w = 5]$, what will the spatial size of the next layer's activation map be if we pad with $P = 2$ and use stride of $S = 2$?

Solution:

$$N^{i+1} = \frac{N^i + F + 2 \cdot P}{S} + 1 = 113$$